



School of Civil Engineering

CE 614 – Statistical and Econometric Methods

Assignment #5 (Discrete Data – FIML Nested Logit)

As in assignments #3 and #4, you are given 151 observations of a travel survey collected in State College Pennsylvania (same data as in assignment #1). All of the households in the sample are making the morning commute to work. They are all departing from the same origin (a large residential complex in the suburbs) and going to work in the Central Business District. They have the choice of three alternate routes; 1) a four-lane arterial (speed limit = 35mph, 2 lanes each direction), 2) a two-lane rural road (speed limit = 35mph, 1 lane each direction) and 3) a limited access four-lane freeway (speed limit = 55mph, 2 lanes each direction).

Your task, using your best model, is to consider all possible nesting structures and comment on your findings with regard to the value of the logsum and the validity of the nesting structure. The structures to test:

1. The freeway by itself with the arterial and rural roads nested.
2. The arterial by itself with the freeway and rural road nested.
3. The rural road by itself with the freeway and arterial nested.

For reference, see pages 334 to 342 in the text. (Washington, S., M. Karlaftis and F. Mannering (2011) Statistical and econometric methods for transportation data analysis, Second Edition, Chapman & Hall/CRC).

Variables available for your specification are (in file LOGIT-A1.txt):

Variable Number	Explanation
x1	Route chosen, rows: 1 - arterial, 2 - rural road, 3 - freeway
x2	Arterial row indicator; 1 for arterial row, 0 for others
x3	Rural row indicator; 1 for rural row, 0 for others
x4	Freeway row indicator; 1 for freeway row, 0 for others
x5	Traffic flow rate
x6	Number of traffic signals
x7	Distance in tenths of miles
x8	Seat belts: 1 - if wear, 0 - if not
x9	Number of passengers in car
x10	Driver age in years: 1 - 18 to 23, 2 - 24 to 29, 3 - 30 to 39, 4 - 40 to 49, 5 - 50 and above
x11	Gender: 1 - male, 0 - female
x12	Marital status: 1 - single, 0 - married
x13	Number of children
x14	Annual income: 1 - less than 20000, 2 - 20000 to 29999, 3 - 30000 to 39999, 4 - 40000 to 49999, 5 - more than 50000
x15	Model year of car (e.g. 86 = 1986)
x16	Origin of car: 1 - domestic, 0 - foreign
x17	Fuel efficiency in miles per gallon

```

--> read;nvar=17;nobs=453;file=D:\old_drive_d\new_laptop\CE697N-disk\LOGIT-A1...
--> create;cage=86-x15$
--> nlogit
;lhs=x1
;choices=arterial,rural,freeway
;tree=nf(rural,arterial),three(freeway)
;model:
u(arterial)=dista*x7/
u(rural)=rural*one+distr*x7/
u(freeway)=freeway*one+distf*x7/
u(nf)=malenf*x12+cagenf*cage+domnf*x16
;ivset: (three)=[1]$

```

```

+-----+
| Discrete choice and multinomial logit models |
+-----+

```

Normal exit from iterations. Exit status=0.

```

+-----+
| FIML Nested Multinomial Logit Model          |
| Maximum Likelihood Estimates                 |
| Model estimated: Oct 03, 2007 at 09:22:15AM. |
| Dependent variable                          | X1 |
| Weighting variable                          | None |
| Number of observations                       | 151 |
| Iterations completed                        | 27 |
| Log likelihood function                      | -94.79945 |
| Number of parameters                        | 9 |
| Info. Criterion: AIC =                      | 1.37483 |
| Finite Sample: AIC =                       | 1.38328 |
| Info. Criterion: BIC =                     | 1.55467 |
| Info. Criterion:HQIC =                    | 1.44789 |
| Restricted log likelihood                   | -198.9332 |
| McFadden Pseudo R-squared                  | .5234610 |
| Chi squared                                 | 208.2676 |
| Degrees of freedom                          | 9 |
| Prob[ChiSqd > value] =                     | .0000000 |
| R2=1-LogL/LogL* Log-L fncn R-sqrd RsqAdj   |
| No coefficients -198.9332 .52346 .50882     |
| Constants only -124.2267 .23688 .21344     |
| At start values -165.8905 .42854 .41099     |
| Response data are given as ind. choice.     |
+-----+

```

```

+-----+-----+-----+-----+-----+
| Variable | Coefficient | Standard Error | b/St.Er. | P[|Z|>z] |
+-----+-----+-----+-----+-----+
+-----+Attributes in the Utility Functions (beta)
DISTA    | -.10342396 | .03211377      | -3.221   | .0013
RURAL    | 2.99039003 | 1.31962302     | 2.266    | .0234
DISTR    | -.15052006 | .03623953     | -4.153   | .0000
FREEWAY  | .68370774  | 3.71037138    | .184     | .8538
DISTF    | -.16064088 | .07611040     | -2.111   | .0348
+-----+Attributes of Branch Choice Equations (alpha)
MALENF   | .79932709  | .68754603     | 1.163    | .2450
CAGENF   | -.14745147 | .06766719     | -2.179   | .0293
DOMNF    | -.96784738 | .73927900     | -1.309   | .1905
+-----+IV parameters, tau(b|l,r),sigma(l|r),phi(r)
NF       | 1.55789246 | .53963034     | 2.887    | .0039
THREE    | 1.00000000 | .....(Fixed Parameter).....

```

```

--> nlogit
;lhs=x1
;choices=arterial,rural,freeway
;tree=na(rural,freeway),three(arterial)
;model:
u(arterial)=artc*one+dista*x7/
u(rural)=rural*one+distr*x7/
u(freeway)=distf*x7/
u(na)=childna*x13+cagena*cage
;ivset: (three)=[1]$

```

```

+-----+
| Discrete choice and multinomial logit models |
+-----+

```

Normal exit from iterations. Exit status=0.

```

+-----+
| FIML Nested Multinomial Logit Model          |
| Maximum Likelihood Estimates                 |
| Model estimated: Oct 03, 2007 at 09:22:15AM. |
| Dependent variable                          | X1 |
| Weighting variable                          | None |
| Number of observations                       | 151 |
| Iterations completed                         | 23 |
| Log likelihood function                      | -95.68433 |
| Number of parameters                         | 8 |
| Info. Criterion: AIC =                       | 1.37330 |
|   Finite Sample: AIC =                       | 1.38002 |
| Info. Criterion: BIC =                       | 1.53316 |
| Info. Criterion:HQIC =                       | 1.43824 |
| Restricted log likelihood                    | -186.4566 |
| McFadden Pseudo R-squared                   | .4868279 |
| Chi squared                                 | 181.5445 |
| Degrees of freedom                          | 8 |
| Prob[ChiSqd > value] =                       | .0000000 |
| R2=1-LogL/LogL*   Log-L fncn   R-sqrd   RsqAdj |
| No coefficients   -186.4566   .48683   .47286 |
| Constants only   -124.2267   .22976   .20880 |
| At start values  -165.8905   .42321   .40751 |
| Response data are given as ind. choice.      |
+-----+

```

```

+-----+-----+-----+-----+-----+
| Variable | Coefficient | Standard Error | b/St.Er. | P[|Z|>z] |
+-----+-----+-----+-----+-----+
-----+Attributes in the Utility Functions (beta)
ARTC      | 1.03855064 | 3.58492590 | .290 | .7720
DISTA     | -.11380007 | .03180433 | -3.578 | .0003
RURAL     | 4.17717602 | 3.38510540 | 1.234 | .2172
DISTR     | -.18662631 | .05521611 | -3.380 | .0007
DISTF     | -.10946256 | .08283669 | -1.321 | .1864
-----+Attributes of Branch Choice Equations (alpha)
CHILDNA   | -.25543085 | .28470512 | -.897 | .3696
CAGENA    | .14335143 | .06753628 | 2.123 | .0338
-----+IV parameters, tau(b|l,r),sigma(l|r),phi(r)
NA        | .87458533 | .38616914 | 2.265 | .0235
THREE     | 1.00000000 | ..... (Fixed Parameter) .....

```

```

--> nlogit
;lhs=x1
;choices=arterial,rural,freeway
;tree=nr(arterial, freeway), three(rural)
;model:
u(arterial)=artc*one+dista*x7/
u(rural)=rural*one+distr*x7/
u(freeway)=distf*x7/
u(nr)=childnr*x13+cagenr*cage
;ivset: (three)=[1]$

```

Normal exit from iterations. Exit status=0.

```

+-----+
| FIML Nested Multinomial Logit Model
| Maximum Likelihood Estimates
| Model estimated: Oct 03, 2007 at 09:31:48AM.
| Dependent variable           X1
| Weighting variable           None
| Number of observations       151
| Iterations completed         19
| Log likelihood function      -97.63937
| Number of parameters         8
| Info. Criterion: AIC =       1.39920
|   Finite Sample: AIC =       1.40591
| Info. Criterion: BIC =       1.55905
| Info. Criterion:HQIC =       1.46414
| Restricted log likelihood    -137.9363
| McFadden Pseudo R-squared   .2921415
| Chi squared                   80.59384
| Degrees of freedom           8
| Prob[ChiSqd > value] =       .0000000
| R2=1-LogL/LogL*   Log-L fncn  R-sqrd  RsqAdj
| No coefficients    -137.9363  .29214  .27288
| Constants only    -124.2267  .21402  .19264
| At start values  -102.4397  .04686  .02092
| Response data are given as ind. choice.
+-----+

```

Variable	Coefficient	Standard Error	b/St.Er.	P[Z >z]
-----+Attributes in the Utility Functions (beta)				
ARTC	2.05344082	2.44765186	.839	.4015
DISTA	-.10053784	.02954438	-3.403	.0007
RURAL	5.81928036	3.44889285	1.687	.0915
DISTR	-.18736641	.03947392	-4.747	.0000
DISTF	-.06492744	.04441577	-1.462	.1438
-----+Attributes of Branch Choice Equations (alpha)				
CHILDNR	.32329418	.26774092	1.207	.2272
CAGENR	-.03904460	.05186271	-.753	.4515
-----+IV parameters, tau(b l,r), sigma(l r), phi(r)				
NR	1.39417533	.55494154	2.512	.0120
THREE	1.00000000(Fixed Parameter).....		